Assignment 8.1

Thip Rattanavilay

2/5/2021

##Dataset 8.1

Fit a Logistic Regression Model to the Thoracic Surgery Binary Dataset

A. Fit a binary logistic regression model to the data set that predicts whether or not the patient survived for one year (the Risk1Y variable) after the surgery. Use the glm() function to perform the logistic regression. See Generalized Linear Models for an example. Include a summary using the summary() function in your results.

```
#Logistic Regression
lgm1 <- glm(Risk1Yr ~ DGN + PRE4 + PRE5 + PRE6 +
            PRE7 + PRE8 + PRE9 + PRE10 + PRE11 +
            PRE14 + PRE17 + PRE19 + PRE25 + PRE30 +
            PRE32 + AGE, data = data, family = binomial()  )
summary(lgm1)
```

```
##
## Call:
## glm(formula = Risk1Yr ~ DGN + PRE4 + PRE5 + PRE6 + PRE7 + PRE8 +
##     PRE9 + PRE10 + PRE11 + PRE14 + PRE17 + PRE19 + PRE25 + PRE30 +
##     PRE32 + AGE, family = binomial(), data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6084  -0.5439  -0.4199  -0.2762   2.4929
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.655e+01  2.400e+03  -0.007  0.99450
## DGNDGN2      1.474e+01  2.400e+03   0.006  0.99510
## DGNDGN3      1.418e+01  2.400e+03   0.006  0.99528
## DGNDGN4      1.461e+01  2.400e+03   0.006  0.99514
## DGNDGN5      1.638e+01  2.400e+03   0.007  0.99455
## DGNDGN6      4.089e-01  2.673e+03   0.000  0.99988
## DGNDGN8      1.803e+01  2.400e+03   0.008  0.99400
## PRE4        -2.272e-01  1.849e-01  -1.229  0.21909
## PRE5        -3.030e-02  1.786e-02  -1.697  0.08971 .
## PRE6PRZ1    -4.427e-01  5.199e-01  -0.852  0.39448
## PRE6PRZ2    -2.937e-01  7.907e-01  -0.371  0.71030
## PRE7T        7.153e-01  5.556e-01   1.288  0.19788
```

```
## PRE8T        1.743e-01  3.892e-01   0.448  0.65419
## PRE9T        1.368e+00  4.868e-01   2.811  0.00494 **
## PRE10T       5.770e-01  4.826e-01   1.196  0.23185
## PRE11T       5.162e-01  3.965e-01   1.302  0.19295
## PRE14OC12    4.394e-01  3.301e-01   1.331  0.18318
## PRE14OC13    1.179e+00  6.165e-01   1.913  0.05580 .
## PRE14OC14    1.653e+00  6.094e-01   2.713  0.00668 **
## PRE17T       9.266e-01  4.445e-01   2.085  0.03709 *
## PRE19T      -1.466e+01  1.654e+03  -0.009  0.99293
## PRE25T      -9.789e-02  1.003e+00  -0.098  0.92227
## PRE30T       1.084e+00  4.990e-01   2.172  0.02984 *
## PRE32T      -1.398e+01  1.645e+03  -0.008  0.99322
## AGE         -9.506e-03  1.810e-02  -0.525  0.59944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 395.61  on 469  degrees of freedom
## Residual deviance: 341.19  on 445  degrees of freedom
## AIC: 391.19
##
## Number of Fisher Scoring iterations: 15
```

B. According to the summary, which variables had the greatest effect on the survival rate?

PRE9T, PRE14OC14, pRE17T, PRE30T all had significant values, hence the greatest effect on survival rate.

C. To compute the accuracy of your model, use the dataset to predict the outcome variable. The percent of correct predictions is the accuracy of your model. What is the accuracy of your model?

88.6%

```
# Split the data into training and validation data sets
split <- sample.split(data, SplitRatio = 0.8)
split
```

```
##  [1]  TRUE FALSE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE  TRUE  TRUE FALSE  TRUE
## [13]  TRUE FALSE  TRUE  TRUE  TRUE
```

```
train <- subset(data, split == "TRUE")
validate <- subset(data, split == "FALSE")
# Train model using training data set
lgm2 <- glm(Risk1Yr ~ DGN + PRE4 + PRE5 + PRE6 +
            PRE7 + PRE8 + PRE9 + PRE10 + PRE11 +
            PRE14 + PRE17 + PRE19 + PRE25 + PRE30 +
            PRE32 + AGE, data = train, family = binomial()  )
summary(lgm2)
```

```
##
## Call:
## glm(formula = Risk1Yr ~ DGN + PRE4 + PRE5 + PRE6 + PRE7 + PRE8 +
##     PRE9 + PRE10 + PRE11 + PRE14 + PRE17 + PRE19 + PRE25 + PRE30 +
##     PRE32 + AGE, family = binomial(), data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5814  -0.5626  -0.4378  -0.2328   2.5896
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -16.45376 2399.54555  -0.007   0.9945
## DGNDGN2       14.94379 2399.54480   0.006   0.9950
## DGNDGN3       14.26240 2399.54476   0.006   0.9953
## DGNDGN4       14.87068 2399.54479   0.006   0.9951
## DGNDGN5       16.98963 2399.54484   0.007   0.9944
## DGNDGN6        0.59562 2673.00472   0.000   0.9998
## DGNDGN8       18.42377 2399.54527   0.008   0.9939
## PRE4          -0.05069    0.33145  -0.153   0.8785
## PRE5          -0.23797    0.35899  -0.663   0.5074
## PRE6PRZ1      -0.44523    0.65817  -0.676   0.4988
## PRE6PRZ2      -0.54728    1.03005  -0.531   0.5952
## PRE7T          1.01471    0.68455   1.482   0.1383
## PRE8T          0.18219    0.47139   0.387   0.6991
## PRE9T          1.23914    0.58936   2.103   0.0355 *
## PRE10T         0.70142    0.60921   1.151   0.2496
## PRE11T         0.28002    0.48911   0.573   0.5670
## PRE14OC12      0.51661    0.37957   1.361   0.1735
## PRE14OC13      0.36418    0.82814   0.440   0.6601
## PRE14OC14      0.34409    0.91984   0.374   0.7084
## PRE17T         0.78496    0.55776   1.407   0.1593
## PRE19T       -14.75875 1667.96169  -0.009   0.9929
## PRE25T        -0.22633    1.25703  -0.180   0.8571
## PRE30T         1.52615    0.73080   2.088   0.0368 *
## PRE32T       -14.63921 2399.54478  -0.006   0.9951
## AGE           -0.02091    0.02244  -0.932   0.3516
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 289.82  on 358  degrees of freedom
## Residual deviance: 249.90  on 334  degrees of freedom
## AIC: 299.9
##
## Number of Fisher Scoring iterations: 15
```

```
# Run validation data through the model built on training data
res <- predict(lgm2, validate, type = "response")
res
```

```
##            2            6           11           14           19           23
## 1.377027e-01 2.726665e-02 6.137973e-02 1.408867e-01 1.315441e-01 1.230180e-01
```

```
##           28           31           36           40           45           48
## 7.935816e-02 1.892930e-01 8.485769e-02 7.129754e-02 2.788355e-01 1.154827e-01
##           53           57           62           65           70           74
## 5.641708e-01 8.758745e-02 1.230651e-01 2.537484e-01 1.470325e-01 1.556198e-02
##           79           82           87           91           96           99
## 1.861082e-01 1.421559e-01 1.123701e-01 1.679714e-01 6.749960e-02 7.192012e-08
##          104          108          113          116          121          125
## 1.109028e-08 1.391332e-01 4.840967e-08 3.390470e-01 3.587940e-02 1.152874e-01
##          130          133          138          142          147          150
## 1.134059e-01 5.236133e-08 1.741567e-01 9.794816e-02 1.299379e-02 5.473398e-02
##          155          159          164          167          172          176
## 9.594121e-02 1.624691e-01 5.461157e-02 1.945705e-01 3.434606e-01 3.916271e-01
##          181          184          189          193          198          201
## 1.471059e-01 1.187176e-01 8.908445e-02 4.633390e-02 3.070363e-02 1.546445e-01
##          206          210          215          218          223          227
## 1.172305e-01 4.062538e-01 8.065222e-02 7.129052e-02 3.021773e-01 1.542133e-01
##          232          235          240          244          249          252
## 7.147261e-01 1.117306e-01 1.273219e-01 3.744062e-02 1.561529e-01 7.683760e-02
##          257          261          266          269          274          278
## 1.025656e-01 1.936019e-02 1.023362e-01 3.636964e-01 2.385587e-01 2.061179e-01
##          283          286          291          295          300          303
## 6.753180e-02 8.363262e-02 1.265220e-01 2.993412e-01 1.202187e-01 6.749673e-02
##          308          312          317          320          325          329
## 1.445449e-01 5.927782e-02 3.164870e-02 9.129358e-09 4.423766e-02 1.711667e-01
##          334          337          342          346          351          354
## 2.641049e-02 1.418398e-01 7.606229e-02 1.862763e-01 1.367746e-01 1.997344e-06
##          359          363          368          371          376          380
## 1.398460e-01 3.531148e-01 7.732797e-01 1.148609e-01 5.517174e-02 1.200479e-01
##          385          388          393          397          402          405
## 3.971697e-02 1.473189e-01 1.601244e-01 6.016293e-02 1.863320e-02 3.451337e-01
##          410          414          419          422          427          431
## 7.282474e-02 1.672479e-01 1.327224e-01 8.645794e-02 1.425310e-01 9.115859e-02
##          436          439          444          448          453          456
## 9.839175e-02 5.123551e-08 1.317247e-02 2.448331e-01 2.371259e-01 2.054748e-01
##          461          465          470
## 4.330110e-02 1.104171e-01 9.102330e-02
```

```
res2 <-predict(lgm2, train, type = "response")
res2
```

```
##            1            3            4            5            7            8
## 2.654470e-01 9.301869e-02 1.434050e-02 1.270730e-01 2.071122e-01 1.100391e-01
##            9           10           12           13           15           16
## 8.009591e-02 1.025202e-01 4.394473e-02 9.412787e-02 7.934725e-02 8.198096e-02
##           17           18           20           21           22           24
## 2.659768e-01 2.158712e-01 5.286760e-02 9.363808e-02 1.570191e-01 6.297503e-02
##           25           26           27           29           30           32
## 4.444827e-01 6.345989e-07 7.172228e-02 1.090949e-01 6.247575e-08 2.768351e-02
##           33           34           35           37           38           39
## 5.773675e-01 7.726608e-02 2.914100e-02 1.240485e-01 1.238890e-01 5.831808e-02
##           41           42           43           44           46           47
## 5.404418e-01 2.155175e-01 1.045871e-01 8.282862e-01 7.308596e-02 9.241607e-02
##           49           50           51           52           54           55
## 1.283763e-01 1.818973e-02 2.498309e-02 6.067642e-02 1.621512e-01 1.063157e-01
```

```
##           56           58           59           60           61           63
## 2.024827e-01 3.761456e-01 1.076961e-01 8.009812e-02 1.102484e-01 5.001978e-02
##           64           66           67           68           69           71
## 3.165771e-02 3.515607e-02 2.653895e-02 2.366594e-01 1.383510e-01 1.229657e-02
##           72           73           75           76           77           78
## 1.989383e-01 5.439505e-02 7.092745e-02 1.467006e-01 1.293603e-01 9.980962e-02
##           80           81           83           84           85           86
## 2.372300e-02 1.216585e-01 1.462956e-01 4.101561e-02 9.865406e-02 1.253409e-01
##           88           89           90           92           93           94
## 1.711851e-01 6.306224e-01 3.154776e-07 7.290829e-02 1.279303e-01 3.956342e-02
##           95           97           98          100          101          102
## 1.222607e-01 2.043550e-01 6.881181e-08 3.564672e-01 6.180880e-02 3.065786e-01
##          103          105          106          107          109          110
## 1.343609e-01 2.571816e-02 1.138215e-01 1.593731e-01 1.245650e-02 2.837128e-01
##          111          112          114          115          117          118
## 1.366292e-01 2.420549e-01 5.637771e-02 1.572444e-01 2.298359e-01 2.945441e-01
##          119          120          122          123          124          126
## 7.731780e-02 1.556690e-01 9.307402e-02 3.909842e-01 9.673278e-02 1.206202e-01
##          127          128          129          131          132          134
## 7.082692e-02 2.636096e-01 3.812829e-01 7.061199e-02 1.217233e-01 8.118840e-02
##          135          136          137          139          140          141
## 1.055914e-01 8.191201e-02 1.953073e-01 1.541152e-01 1.878187e-02 1.786684e-01
##          143          144          145          146          148          149
## 7.358750e-03 1.813758e-01 1.433527e-01 1.138647e-01 1.207307e-01 9.273231e-02
##          151          152          153          154          156          157
## 2.902038e-02 7.733679e-02 3.497828e-02 1.031937e-01 4.234472e-02 2.839971e-01
##          158          160          161          162          163          165
## 9.723657e-08 1.054955e-01 2.065422e-02 7.286373e-02 3.098214e-01 2.109421e-01
##          166          168          169          170          171          173
## 4.284707e-01 8.612320e-02 1.595597e-01 1.198241e-01 1.273790e-01 4.796265e-01
##          174          175          177          178          179          180
## 9.515792e-02 1.486009e-01 1.126893e-01 1.379979e-01 1.570071e-01 1.567712e-01
##          182          183          185          186          187          188
## 8.672847e-02 7.807777e-02 1.752243e-02 4.528034e-01 8.114500e-02 1.606500e-01
##          190          191          192          194          195          196
## 1.268303e-01 6.389220e-08 6.745145e-02 9.527902e-02 6.166054e-02 1.643859e-01
##          197          199          200          202          203          204
## 1.830962e-01 3.102842e-02 1.765564e-01 8.212173e-02 3.072781e-01 1.476386e-01
##          205          207          208          209          211          212
## 1.578568e-02 6.048879e-02 8.490661e-02 6.940278e-02 4.496835e-02 1.284610e-01
##          213          214          216          217          219          220
## 3.740002e-01 2.465949e-01 7.756204e-02 2.501395e-01 5.577649e-02 5.974086e-02
##          221          222          224          225          226          228
## 7.050023e-01 1.448667e-01 4.226649e-02 1.240614e-01 4.660690e-01 1.274247e-01
##          229          230          231          233          234          236
## 1.727214e-02 2.506580e-01 2.201113e-01 8.250197e-02 1.528190e-01 9.400989e-02
##          237          238          239          241          242          243
## 1.480547e-01 9.986329e-02 5.199163e-01 3.604854e-02 8.648462e-02 1.994060e-01
##          245          246          247          248          250          251
## 3.499449e-08 1.142379e-01 9.349010e-02 1.728873e-01 1.216796e-01 1.017732e-01
##          253          254          255          256          258          259
## 1.088820e-01 1.269774e-01 8.287617e-02 1.279501e-08 7.626528e-02 1.000053e-01
##          260          262          263          264          265          267
## 1.087573e-01 1.471164e-01 1.715944e-01 2.011617e-02 8.821210e-02 9.397337e-02
```

```
##          268          270          271          272          273          275
## 1.143475e-01 1.328385e-01 1.633214e-01 4.727154e-01 2.691042e-02 2.275142e-01
##          276          277          279          280          281          282
## 1.108123e-01 1.264598e-01 1.207882e-02 9.655314e-02 1.022945e-01 1.695403e-02
##          284          285          287          288          289          290
## 2.923151e-01 7.776321e-02 1.645628e-01 1.320413e-01 1.629201e-01 1.016112e-01
##          292          293          294          296          297          298
## 1.566315e-01 6.389220e-08 7.605334e-02 1.190128e-01 1.671745e-01 5.598399e-01
##          299          301          302          304          305          306
## 7.413546e-02 1.536191e-01 2.704941e-02 1.346629e-01 6.917873e-02 1.274116e-01
##          307          309          310          311          313          314
## 7.136272e-01 9.378481e-02 5.146044e-02 2.259135e-02 2.954703e-01 1.272550e-01
##          315          316          318          319          321          322
## 2.315619e-01 2.633411e-01 2.861604e-01 8.826183e-02 9.106433e-02 6.677221e-02
##          323          324          326          327          328          330
## 7.969826e-02 3.342656e-01 1.013521e-09 1.254078e-01 1.094001e-01 3.843884e-02
##          331          332          333          335          336          338
## 6.667009e-09 5.097732e-02 7.675150e-02 1.740906e-01 9.137427e-02 1.810044e-01
##          339          340          341          343          344          345
## 5.026558e-02 1.478187e-01 4.938698e-02 1.560205e-01 1.077425e-01 1.036476e-01
##          347          348          349          350          352          353
## 9.852067e-02 2.153383e-01 1.156872e-01 9.317764e-11 8.815602e-02 3.247222e-09
##          355          356          357          358          360          361
## 6.192239e-02 1.202301e-01 3.715317e-01 1.153725e-01 4.509363e-02 1.572941e-01
##          362          364          365          366          367          369
## 1.044740e-01 1.921796e-01 1.268981e-01 1.172945e-01 8.639587e-02 1.110976e-07
##          370          372          373          374          375          377
## 8.894156e-02 3.353737e-02 1.100910e-01 4.085462e-01 1.223288e-01 7.030293e-02
##          378          379          381          382          383          384
## 1.168588e-01 6.540347e-02 1.851186e-01 5.175188e-02 1.517027e-01 2.628292e-02
##          386          387          389          390          391          392
## 2.302668e-01 2.289252e-01 1.278807e-01 2.288171e-01 1.042467e-01 3.458464e-01
##          394          395          396          398          399          400
## 1.120353e-01 2.275982e-01 2.004387e-01 9.921156e-02 1.259412e-01 8.652064e-02
##          401          403          404          406          407          408
## 1.846833e-02 1.578950e-01 1.578911e-01 2.282222e-08 7.951397e-02 1.823169e-01
##          409          411          412          413          415          416
## 2.630742e-01 2.559743e-01 1.937233e-01 1.582767e-02 1.234265e-01 1.407059e-02
##          417          418          420          421          423          424
## 1.545625e-01 4.922023e-02 2.261791e-01 4.453953e-01 9.433253e-02 3.695787e-02
##          425          426          428          429          430          432
## 2.941113e-01 1.563265e-01 4.368752e-02 1.720494e-01 1.734673e-01 1.213088e-01
##          433          434          435          437          438          440
## 3.237181e-02 1.890637e-01 9.117891e-02 3.067126e-01 1.548647e-01 1.208295e-01
##          441          442          443          445          446          447
## 1.539491e-01 3.290176e-02 1.804753e-01 8.067507e-08 7.317209e-02 5.555173e-01
##          449          450          451          452          454          455
## 9.328487e-02 1.691391e-01 2.018765e-02 2.256128e-01 1.045141e-01 4.975437e-02
##          457          458          459          460          462          463
## 1.672442e-01 9.694226e-02 2.202693e-02 3.482299e-02 1.005288e-01 1.120620e-01
##          464          466          467          468          469
## 4.752355e-01 2.007421e-01 5.347921e-02 3.532662e-02 1.464074e-01
```

```r
#Validate model using confusion matrix
confmatrix <- table(Actual_Value=train$Risk1Yr, Predicted_Value = res2 >0.5)
confmatrix
```

```
##              Predicted_Value
## Actual_Value FALSE TRUE
##            F   305    4
##            T    45    5
```

```r
#Accuracy
(confmatrix[[1,1]] + confmatrix[[2,2]])/sum(confmatrix)
```

```
## [1] 0.8635097
```